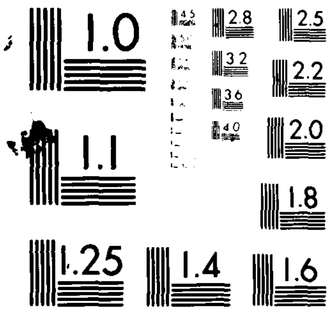
AD-A111 962 ROYAL SIGNALS AND RADAR ESTABLISHMENT MALVERN (ENGLAND) F/G 9/1
THE CALCULATION OF THE EQUIVALENT CIRCUITS OF COAXIAL-LINE STEP--ETC(U)
1981 T E HODGETTS
UNCLASSIFIED RSRE-MEMO-3422 DRIC-BR-81999 NL

1 of 1
AD
A111962

END
DATE
FILMED
4 82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963 A

RSRE
MEMORANDUM No. 3422

ROYAL SIGNALS & RADAR ESTABLISHMENT

THE CALCULATION OF THE EQUIVALENT CIRCUITS OF COAXIAL-LINE STEP DISCONTINUITIES

Author: T E Hodgetts

PROCUREMENT EXECUTIVE,
MINISTRY OF DEFENCE,
RSRE MALVERN,
WORCS.

RSRE MEMORANDUM No. 3422

ROYAL SIGNALS AND RADAR ESTABLISHMENT

Memorandum 3422

Title: THE CALCULATION OF THE EQUIVALENT CIRCUITS OF COAXIAL-LINE STEP DISCONTINUITIES

Author: T E Hodgetts

SUMMARY

This memorandum describes a method of calculating the equivalent circuits of discontinuities in coaxial lines caused by abrupt changes of radius in one or both of the coaxial conductors. The theory is fully worked out for discontinuities of the inner conductor alone; the implementation of the final equations on a large computer is described; and the performance of the resulting program is discussed, including the effects of truncation error and extrapolation to infinity. Finally, the use of the theory in connection with coaxial-line reflection and transmission standards is described, and a simple error analysis of these standards is presented, which shows that (at present) their performance is calculable to an accuracy limited only by the precision to which they can be made.

This memorandum is for advance information. It is not necessarily to be regarded as a final or official statement by Procurement Executive, Ministry of Defence

THE CALCULATION OF THE EQUIVALENT CIRCUITS OF COAXIAL-LINE STEP DISCONTINUITIES

T E Hodgetts

LIST OF CONTENTS

1 INTRODUCTION

This memorandum describes a practical method of calculating the equivalent circuits of coaxial-line discontinuities of the types shown in Figs 1a-1e. The method used can be applied to many similar problems in one-conductor or two-conductor waveguide systems, as described by Harrington[1]; but Harrington's treatment is completely general and gives no feeling for the computational details which are essential when the method is applied to a particular case. Accordingly, the treatment presented here discusses fully two particular cases which are important in practice, and uses them to illustrate the general theory, thus complementing Harrington's exposition.

The method applies in principle to any problem involving semi-infinite "generalised cylindrical waveguides" joined at a plane at right-angles to their common axis. (A generalised cylindrical waveguide is defined by translating a simple closed plane curve at right-angles to its own plane, as shown in Figs 2.) The behaviour of the electromagnetic field in the neighbourhood of the discontinuity plane in such a system can be characterised by an equivalent circuit located at that plane; the discussion which follows will show this, establish that the form of the equivalent circuit is a single shunt element (in general) and obtain an expression for its value. The numerical aspects of calculating the value will be described also, as will the practical importance of the work (for standards of coaxial-line impedance) and the factors affecting accuracy.

2 PRELIMINARIES ON FIELDS AND MODES

The first stage in the treatment of any particular case is to express the fields in the guides in terms of the permitted modes for that case. The theory of modes in generalised cylinders has been fully discussed elsewhere (eg [1], [2] and [3]), so this section will only present a summary of the important facts.

If a region of space (such as the interior of a generalised cylinder) is linear, isotropic, homogeneous and non-conducting, and contains no free charges or current-carrying elements, any electromagnetic field within it may be written as the sum of two partial fields, one of which has its $\underline{E}$ vector transverse to an arbitrary fixed direction in space (a transverse electric or TE field) while the other has its $\underline{H}$ vector transverse to the same arbitrary fixed direction (a transverse magnetic or TM field); and each partial field can be derived from a single scalar function which satisfies the wave equation. When we deal with

generalised cylinders having a common axis it is convenient to make that axis the arbitrary fixed direction and call it the z-direction; then fields which vary in time as $\exp(j\omega t)$ can be derived from scalar functions Ψ satisfying Helmholtz' equation

$$\left(\nabla_T^2 + \frac{\partial^2}{\partial z^2} + k^2\right)\Psi = 0 \tag{1}$$

where k is the free-medium angular wavenumber corresponding to ω and ∇_T^2 is the two-dimensional Laplacian operator transverse to z. These functions Ψ satisfy boundary conditions which are independent of z, and so eqn (1) is separable, giving

$$(\nabla_T^2 + k'^2)\psi = 0 \tag{2a}$$

with

$$\Psi = \psi\zeta \quad , \tag{2b}$$

$$\frac{d^2\zeta}{dz^2} - \gamma^2\zeta = 0 \tag{2c}$$

and

$$\gamma^2 = k'^2 - k^2 \quad . \tag{2d}$$

The boundary conditions on Ψ become boundary conditions on ψ; and it turns out that eqn (2a) admits non-trivial solutions only for particular non-negative real values of k'^2, and corresponding values of γ^2 from eqn (2d), each with its own corresponding field. Each of these sets (k'^2, γ^2 and corresponding field) constitutes a mode, and any possible field within a generalised cylinder can be expressed as a linear combination of all the modes for that cylinder. (This does not conflict with the division into TE and TM fields, because each mode can be classified as TE or TM.)

Power is freely propagated only by the fields of those modes which admit oscillatory, rather than exponentially-decaying, solutions to eqn (2c), that is, those which have negative values of γ^2. There are only a finite number of these, and the number depends on the frequency; in one-conductor systems at sufficiently

low frequencies there are none at all. However, in two-conductor systems there is always one mode having $k'^2 = 0$, which will propagate power at any frequency. This mode belongs to the TM class, but has special properties and is often considered to be in a class of its own; because its $\underline{E}$ and $\underline{H}$ vectors are both transverse to the z-axis, it is called transverse electromagnetic (TEM).

3 ROTATIONALLY-SYMMETRIC SYSTEMS

It is possible to write down the modes which can exist in systems like those in Fig 1, from the discussions in [1] and [2]. Alternatively, we can derive them directly by using rotational symmetry about the z-axis instead of the translational symmetry along it.

Subject to the restrictions introduced already, Maxwell's equations for a field varying as $\exp(j\omega t)$ take the form (see [2])

$$\text{curl } \underline{E} + j\omega\mu\underline{H} = \underline{0} \qquad (3a) \qquad\qquad \text{div } \underline{H} = 0 \qquad (3b)$$

$$\text{curl } \underline{H} - j\omega\varepsilon\underline{E} = \underline{0} \qquad (3c) \qquad\qquad \text{div } \underline{E} = 0 \qquad (3d)$$

in the usual notation.

We introduce cylindrical polar co-ordinates, expand the vector operators in eqns (3) and enforce rotational symmetry by setting to zero all the partial derivatives with respect to ϕ, giving (see Fig 3)

$$-\frac{\partial E_\phi}{\partial z} + j\omega\mu H_\rho = 0 \qquad (4a)$$

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho E_\phi) + j\omega\mu H_z = 0 \qquad (4b)$$

$$\frac{\partial H_\rho}{\partial z} - \frac{\partial H_z}{\partial \rho} - j\omega\varepsilon E_\phi = 0 \qquad (4c)$$

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho H_\rho) + \frac{\partial H_z}{\partial z} = 0 \qquad (4d)$$

$$-\frac{\partial H_\phi}{\partial z} - j\omega\varepsilon E_\rho = 0 \qquad (5a)$$

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho H_\phi) - j\omega\varepsilon E_z = 0 \qquad (5b)$$

$$\frac{\partial E_\rho}{\partial z} - \frac{\partial E_z}{\partial \rho} + j\omega\mu H_\phi = 0 \tag{5c}$$

$$\frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho E_\rho) + \frac{\partial E_z}{\partial \rho} = 0 \tag{5d}$$

On substituting eqns (4a) and (4b) into eqns (4c) and (4d), we find that (4d) is satisfied identically and that (4c) takes the form

$$\frac{\partial^2}{\partial \rho^2}(\rho E_\phi) - \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho E_\phi) + \frac{\partial^2}{\partial z^2}(\rho E_\phi) + k^2(\rho E_\phi) = 0 \tag{6}$$

with

$$k^2 = \omega^2\mu\varepsilon \tag{6a}$$

as usual. A similar treatment of eqns (5) leads to

$$\frac{\partial^2}{\partial \rho^2}(\rho H_\phi) - \frac{1}{\rho}\frac{\partial}{\partial \rho}(\rho H_\phi) + \frac{\partial^2}{\partial z^2}(\rho H_\phi) + k^2(\rho H_\phi) = 0 \tag{7}$$

The boundary conditions to be used with eqns (6) and (7) follow from the requirements that the tangential electric field on any perfectly-conducting surface must vanish. Using eqns (5a) and (5b), we deduce

$$(\rho E_\phi) = 0 \qquad (8a) \qquad\qquad \frac{\partial}{\partial n}(\rho H_\phi) = 0 \qquad (8b)$$

where n denotes a normal to the boundary surface (or surfaces). Eqns (4), (6) and (8a) specify a TE field, E_ϕ being obtained by solving (6) subject to (8a) and the dependent components H_ρ and H_z being given by (4a) and (4b); similarly eqns (5), (7) and (8b) specify a TM field having components H_ϕ, E_ρ and E_z.

4 CYLINDRICAL SYMMETRY

To apply these results to the systems represented in Figs 1, we begin by solving eqn (7) using eqns (2). Writing

$$(\rho H_\phi) = \psi(\rho)\zeta(z) \tag{9a}$$

we obtain

$$\frac{d^2\zeta}{dz^2} - \gamma^2\zeta = 0 \quad \text{and} \quad \gamma^2 = k'^2 - k^2, \text{ as in (2), and}$$

$$\frac{d^2\psi}{d\rho^2} - \frac{1}{\rho}\frac{d\psi}{d\rho} + k'^2\psi = 0 \quad . \tag{9b}$$

If $k'^2 = 0$, eqn (9b) has the solution $\psi = A\rho^2 + B$, where A and B are arbitrary constants; otherwise the solution is

$$\psi = (k'\rho)(A'J_1(k'\rho) + B'Y_1(k'\rho)) \tag{9c}$$

where A' and B' are arbitrary constants and J_1 and Y_1 are Bessel's functions of the first and second kinds of order 1, as may be shown by substitution. Eqn (6) gives similar solutions for (ρE_ϕ), but we shall see that these are not actually needed.

The systems of Figs 1 are made up of right circular cylinders butted together. Accordingly, everywhere except at the plane of discontinuity the boundary conditions (eqns (8)) take the form $(\rho E_\phi) = 0$ and $\frac{\partial}{\partial\rho}(\rho H_\phi) = 0$ at one or two particular values of ρ, depending on whether the system contains one or two conductors; and if there is only one conductor the field components must be finite at $\rho = 0$.

It is usual to operate coaxial systems under conditions such that the TEM mode is the only one which can freely propagate power; from the boundary conditions and the special solution we find that this mode is given by $(\rho H_\phi) = \psi\zeta = B\zeta$ (no non-trivial special solution for (ρE_ϕ) is possible, and the solution for (ρH_ϕ) also vanishes if there is only one conductor). The TEM mode is a TM field, and systems like those of Figs 1 (with rotationally-symmetrical discontinuities as well as conductors) have independent TM and TE fields. It follows that only a TM field can exist in the systems of Figs 1 when they are worked so that the TEM mode is the only one freely propagating; any TE field travelling inwards from infinity would decay to zero before encountering the plane of discontinuity. Consequently, if we obtain the permitted modes for the systems of Figs 1 by treating them as generalised cylinders, we must reject all TE modes, as well as all TM modes which are not rotationally-symmetrical.

If these rejected modes are considered, it turns out that at least one of them begins to propagate power freely at a frequency at which all the retained modes are still cut off. (See [2] for a full discussion.) Because of this, there are two "critical frequencies" for any system of the type considered here; a lower critical frequency above which at least one (TE) mode propagates freely, and an upper critical frequency above which at least one of the retained modes propagates freely. (Conveniently, the TM mode with the lowest cut-off frequency is one of the retained modes, otherwise - since the TE modes and the rotationally-asymmetrical modes are rejected for different reasons - there would be three critical frequencies.) The analysis which follows is formally valid below the upper critical frequency, but will only hold between the two critical frequencies provided that all TE modes have been filtered out before one of the discontinuities represented in Figs 1 is encountered. Expressions for the critical frequencies, after [2], will be given later.

5 DEVELOPMENT OF THE PROPER MODAL EXPANSION

We are now ready to construct the modal expansion of the allowed field in any of the systems of Figs 1; the one shown in Fig 1b will be used, because its practical importance was the reason for undertaking the work described here. In this case the boundary condition

$$\frac{\partial}{\partial \rho}(\rho H_\phi) = 0 = \frac{d\psi}{d\rho}$$

must be satisfied at $\rho = r_A$ and $\rho = R$ on the A-side of the discontinuity plane, and at $\rho = r_B$ and $\rho = R$ on the B-side. The TEM mode already satisfies this condition; for the TM modes, using eqn (9c) and the relations

$$\frac{d}{dx}(xJ_1(x)) = xJ_o(x) \text{ and } \frac{d}{dx}(xY_1(x)) = xY_o(x) \quad ,$$

we obtain

$$(k'\rho)(A'J_o(k'\rho) + B'Y_o(k'\rho)) = 0 \tag{10}$$

to be satisfied at the proper values of ρ.

On the A-side, the only non-trivial values of k' satisfying eqn (10) are those which satisfy

$$J_o(k'r_A)Y_o(k'R) - J_o(k'R)Y_o(k'r_A) = 0 \tag{10a}$$

From Watson [4], the roots of eqn (10a) are real, positive and distinct; they will be denoted by k_{Ai}, where i is a positive integer and $k_{A1} < k_{A2} < k_{A3} \ldots$. Similarly, on the B-side we have a set of k_{Bi} ($k_{B1} < k_{B2} < k_{B3} \ldots$) satisfying

$$J_o(k'r_B)Y_o(k'R) - J_o(k'R)Y_o(k'r_B) = 0 \tag{10b}$$

The boundary condition also determines, for each allowed k', the proportions of the combination of $J_1(k'\rho)$ and $Y_1(k'\rho)$ in the corresponding ψ. We must take care that this proportionality does not become infinite, so we write the allowed ψ functions on the A-side as

$$\psi_{Ai} = A'_{Ai}(k_{Ai}\rho)\left(J_1(k_{Ai}\rho) + \left(\frac{B'}{A'}\right)_{Ai} Y_1(k_{Ai}\rho)\right) \tag{11}$$

with a similar equation on the B-side; the degenerate form of this is then $A'_{Ai}(k_{Ai}\rho)(J_1(k_{Ai}\rho))$ which is the proper function for the degenerate form of Fig 1b (namely Fig 1a). The boundary condition $d\psi_{Ai}/d\rho = 0$ at $\rho = r_A$ and $\rho = R$ then gives, taking $A'_{Ai} \neq 0$,

$$\left(\frac{B'}{A'}\right)_{Ai} = -\frac{J_o(k_{Ai}r_A)}{Y_o(k_{Ai}r_A)} = -\frac{J_o(k_{Ai}R)}{Y_o(k_{Ai}R)} \tag{12}$$

We see that eqn (10a) is the consistency condition associated with eqn (12). A similar result holds good on the B-side.

We now define the following six families of functions:

$$\left.\begin{aligned} Z_o(k_{Ai}\rho) &= J_o(k_{Ai}\rho) + (B'/A')_{Ai}Y_o(k_{Ai}\rho) \\ Z_1(k_{Ai}\rho) &= J_1(k_{Ai}\rho) + (B'/A')_{Ai}Y_1(k_{Ai}\rho) \\ Z_2(k_{Ai}\rho) &= J_2(k_{Ai}\rho) + (B'/A')_{Ai}Y_2(k_{Ai}\rho) \end{aligned}\right\} \tag{13a}$$

$$\left.\begin{aligned} Z_o(k_{Bi}\rho) &= J_o(k_{Bi}\rho) + (B'/A')_{Bi}Y_o(k_{Bi}\rho) \\ Z_1(k_{Bi}\rho) &= J_1(k_{Bi}\rho) + (B'/A')_{Bi}Y_1(k_{Bi}\rho) \\ Z_2(k_{Bi}\rho) &= J_2(k_{Bi}\rho) + (B'/A')_{Bi}Y_2(k_{Bi}\rho) \end{aligned}\right\} \tag{13b}$$

noting that these Z functions must always have arguments of the form $(k_{Ai}\rho)$ or $(k_{Bi}\rho)$ so that the (B'/A') factor is defined. From eqns (11), (12) and (13) we obtain

$$Z_o(k_{Ai}r_A) = Z_o(k_{Ai}R) = Z_o(k_{Bi}r_B) = Z_o(k_{Bi}P) = 0 \tag{14}$$

for all i, conveniently expressing the boundary conditions, and

$$\psi_{Ai} = A'_{Ai}(k_{Ai}\rho)Z_1(k_{Ai}\rho) \tag{15a}$$

with a similar equation on the B-side.

To complete the construction of the mode functions, we refer to eqn (9a) and the equations immediately following it; these give

$$(\rho H_\phi)_{Ai} = \psi_{Ai}(\rho)\zeta_{Ai}(z) \tag{15b}$$

with $\zeta_{Ai}(z)$ satisfying

$$\frac{d^2\zeta_{Ai}}{dz^2} - \gamma_{Ai}^2\zeta_{Ai} = 0$$

and

$$\gamma_{Ai}^2 = k_{Ai}^2 - k_A^2 \tag{15c}$$

From eqn (6a) we have

$$k_A^2 = \omega^2\mu_A\varepsilon_A \tag{15d}$$

in an obvious notation; and, as always, similar equations hold on the B-side. The functions $\zeta(z)$ for the modes which do not propagate power freely must give fields which decay as we move away from the discontinuity plane; so, taking z to be zero at that plane and positive on the B-side, we have ζ functions of the form $e^{\gamma_{Ai}z}$ and $e^{-\gamma_{Bi}z}$ (taking γ_{Ai}, γ_{Bi} all positive). No arbitrary multiplying constant is needed with these functions, since there is already one in each ψ and only the products $\psi\zeta$ are physically significant. For the TEM mode we must include fields freely propagating in both directions and on both sides of the

discontinuity plane, and from eqn (15c) we have $\gamma_{TEM} = \pm jk_A$ or $\pm jk_B$. The complete field is a linear combination of all these possible mode fields (by [1] and [2]), so, if we write

$$
\left.\begin{aligned}
A'_{Ai} &= -\frac{j\omega\varepsilon_A}{k_{Ai}\gamma_{Ai}}\alpha_i & \qquad \eta_A &= \sqrt{\frac{\mu_A}{\varepsilon_A}} \\
A'_{Bi} &= \frac{j\omega\varepsilon_B}{k_{Bi}\gamma_{Bi}}\beta_i & \qquad \eta_B &= \sqrt{\frac{\mu_B}{\varepsilon_B}}
\end{aligned}\right\} \tag{16}
$$

in terms of the material constants μ_A, ε_A, μ_B and ε_B, then eqns (15) give

$$
\left.\begin{aligned}
H_\phi\ (z<0) &= \frac{1}{\eta_A\rho}\alpha_o\left(e^{-jk_Az} - \Gamma_A e^{jk_Az}\right) + \\
&\quad + \sum_{i=1}^{\infty}\left(-\frac{j\omega\varepsilon_A}{\gamma_{Ai}}\right)\alpha_i e^{\gamma_{Ai}z}\ Z_1(k_{Ai}\rho) \\
H_\phi\ (z>0) &= \frac{1}{\eta_B\rho}\beta_o\left(-e^{jk_Bz} + \Gamma_B e^{-jk_Bz}\right) + \\
&\quad + \sum_{j=1}^{\infty}\left(\frac{j\omega\varepsilon_B}{\gamma_{Bj}}\right)\beta_j e^{-\gamma_{Bj}z}\ Z_1(k_{Bj}\rho)
\end{aligned}\right\} \tag{17a}
$$

The signs and multiplying factors here have been chosen for convenience later. These equations hold for the degenerate case, Fig 1a, as well as in the case of Fig 1b, except that β_o vanishes when the inner conductor is truncated. To complete the representation, we use eqns (5) and (17a) to obtain

$$
\left.\begin{aligned}
E_\rho\ (z<0) &= \frac{1}{\rho}\alpha_o\left(e^{-jk_Az} + \Gamma_A e^{jk_Az}\right) + \sum_{i=1}^{\infty}\alpha_i e^{\gamma_{Ai}z}\ Z_1(k_{Ai}\rho) \\
E_\rho\ (z>0) &= \frac{1}{\rho}\beta_o\left(e^{jk_Bz} + \Gamma_B e^{-jk_Bz}\right) + \sum_{j=1}^{\infty}\beta_j e^{-\gamma_{Bj}z}\ Z_1(k_{Bj}\rho)
\end{aligned}\right\} \tag{17b}
$$

and

$$\left.\begin{aligned} E_z \ (z < 0) \ &= \sum_{i=1}^{\infty} \left(- \frac{k_{Ai}}{\gamma_{Ai}}\right) \alpha_i e^{\gamma_{Ai} z} \ Z_o(k_{Ai}\rho) \\ E_z \ (z > 0) \ &= \sum_{j=1}^{\infty} \left(\frac{k_{Bj}}{\gamma_{Bj}}\right) \beta_j e^{-\gamma_{Bj} z} \ Z_o(k_{Bj}\rho) \end{aligned}\right\} \qquad (17c)$$

To obtain eqns (17c) we use, as before, the relations

$$\frac{d}{dx}\,(xJ_1(x)) \ = \ xJ_o(x) \quad \text{and} \quad \frac{d}{dx}\,(xY_1(x)) \ = \ xY_o(x)$$

together with eqns (13) to deduce that

$$\frac{d}{d(k_{Ai}\rho)}\,((k_{Ai}\rho)Z_1(k_{Ai}\rho)) \ = \ (k_{Ai}\rho)Z_o(k_{Ai}\rho)$$

with a similar equation in k_{Bj}. Eqns (14) and (17c) confirm that our representation satisfies all the boundary conditions introduced so far.

Γ_A and Γ_B can be regarded as voltage reflection coefficients if E_ρ, or some quantity proportional to it, is taken as the measure of "voltage". We will refer to this later.

6 GENERALISED MODES AND CRITICAL FREQUENCIES

The most general set of modes in coaxial systems includes others not derived here, depending on Bessel functions of higher orders and exhibiting rotational asymmetry. These are of no significance in the theory of the equivalent circuit of the discontinuity, but they must be taken into account in determining the critical frequencies referred to previously.

The upper critical frequency is derived from eqns (10), since it turns out that the TM mode with the lowest frequency of free propagation is rotationally-symmetrical (it is the one usually denoted TM_{01}); it begins to propagate freely at the lower of the two frequencies ν_A, ν_B given by

$$2\pi \nu_A \sqrt{\mu_A \varepsilon_A} \ \equiv \ k_A \ = \ k_{A1} \qquad (18a)$$

$$2\pi\nu_B\sqrt{\mu_B\varepsilon_B} \equiv k_B = k_{B1} \tag{18b}$$

using the notation already introduced. The lower critical frequency, however, relates to a rotationally-asymmetrical mode, the so-called TE_{11} mode (which is actually a pair of degenerate modes distinguished only by their planes of polarisation); and to characterise this mode-pair we must replace eqns (10) by

$$J_1'(k'r_A)Y_1'(k'R) - J_1'(k'R)Y_1'(k'r_A) = 0 \tag{19a}$$

and

$$J_1'(k'r_B)Y_1'(k'R) - J_1'(k'R)Y_1'(k'r_B) = 0 \tag{19b}$$

where the prime on the Bessel functions denotes differentiation. The lowest values of k' satisfying these equations are then used in eqns (18) to find the critical frequency.

Good approximate solutions of these equations may be deduced from the discussion in [2] and are as follows:

$$\text{lower critical frequency} = \text{smaller of } 1/(\pi\sqrt{\mu_A\varepsilon_A}\,(R + r_A))$$

$$\text{and } 1/(\pi\sqrt{\mu_B\varepsilon_B}\,(R + r_B))$$

$$\text{upper critical frequency} = \text{smaller of } 1/(2\sqrt{\mu_A\varepsilon_A}\,(R - r_A))$$

$$\text{and } 1/(2\sqrt{\mu_B\varepsilon_B}\,(R - r_B)) \quad .$$

If we take R = 3.5 mm, $r_A = r_B = 1.52$ mm, $\mu_A = \mu_B = \mu_o$, $\varepsilon_A = \varepsilon_B = \varepsilon_o$ (the parameters of a 7 mm-bore precision air-spaced coaxial line), these approximate formulae give the critical frequencies as 19.0 and 75.7 GHz, the exact values being 19.4 and 75.1 GHz to the same number of significant figures. It is worth noting that changing the radius of the inner conductor causes the critical frequencies to move in opposite directions, so, when $\mu_A = \mu_B$ and $\varepsilon_A = \varepsilon_B$ (as is usually the case) the lower critical frequency is fixed by r_A and the upper one by r_B. (r_A is greater than r_B; see Fig 1b.) These conclusions need modifications in detail for the systems of Figs 1c and 1d, but the principles are the same.

For the system of Fig 1a, eqns (10b) and (19b) reduce to the much simpler forms

$$J_o(k'R) = 0 \qquad (20a) \qquad \text{and} \qquad J_1'(k'R) = 0 \qquad (20b)$$

The lowest values of k' satisfying these equations are very nearly (2.405/R) and (1.841/R) respectively, from which we obtain approximate critical frequencies as follows:

$$\text{lower critical frequency} = \text{smaller of } 1/(\pi\sqrt{\mu_A\varepsilon_A}\,(R + r))$$

$$\text{and } 0.920/(\pi\sqrt{\mu_B\varepsilon_B}\,R)$$

$$\text{upper critical frequency} = \text{smaller of } 1/(2\sqrt{\mu_A\varepsilon_A}\,(R - r))$$

$$\text{and } 1.202/(\pi\sqrt{\mu_B\varepsilon_B}\,R) \quad .$$

If we take the same dimensions and material constants as before, the approximate critical frequencies are 19.0 and 32.8 GHz, the exact values being 19.4 and 32.8 GHz to the same number of significant figures. The interesting point here is the large reduction in the upper critical frequency; this makes the behaviour of the equivalent circuit change with frequency considerably more rapidly than it does in the ordinary coaxial system (Fig 1b).

7 THE FIELDS IN THE PLANE OF DISCONTINUITY

We are now ready to begin the derivation proper. In the introduction it was promised that this would be presented in a way which illustrated the proper approach to other similar problems; so we note that equations corresponding to eqns (17) will exist for any problem involving generalised cylinders and the first stage of any derivation is always to find the corresponding equations representing the fields. (Various important special cases are discussed in [2].) Next, we determine the transverse fields at the discontinuity plane, and apply the boundary condition that they must be continuous (since the tangential components of $\underline{E}$ and $\underline{H}$ are continuous across any surface not supporting a surface current). There is one other boundary condition to be applied: the transverse components of $\underline{E}$ must vanish on those portions of the discontinuity plane which are parts of the perfectly-conducting cylinders ($r_B \leq \rho \leq r_A$ in the system of Fig 1b, $0 \leq \rho \leq r$ in that of Fig 1a, for example). All this is readily

generalised; in our particular case, writing $\mathcal{E}_\rho$ and $\mathcal{H}_\phi$ for the values of E_ρ and H_ϕ at the discontinuity plane $z = 0$, we obtain

$$\left.\begin{aligned}
\mathcal{E}_\rho &= \frac{1}{\rho}\,\alpha_o(1 + \Gamma_A) + \sum_{i=1}^{\infty} \alpha_i Z_1(k_{Ai}\rho) && (r_A < \rho < R) \\
&= \frac{1}{\rho}\,\beta_o(1 + \Gamma_B) + \sum_{j=1}^{\infty} \beta_j Z_1(k_{Bj}\rho) && (r_B < \rho < R) \\
&= 0 && (r_B < \rho < r_A)
\end{aligned}\right\} \quad (21)$$

$$\left.\begin{aligned}
\mathcal{H}_\phi &= \frac{1}{\eta_A\rho}\,\alpha_o(1 - \Gamma_A) + \sum_{i=1}^{\infty}\left(-\frac{j\omega\varepsilon_A}{\gamma_{Ai}}\right)\alpha_i Z_1(k_{Ai}\rho) && (r_A < \rho < R) \\
&= \frac{1}{\eta_B\rho}\,\beta_o(-1 + \Gamma_B) + \sum_{j=1}^{\infty}\left(\frac{j\omega\varepsilon_B}{\gamma_{Bj}}\right)\beta_j Z_1(k_{Bj}\rho) && (r_B < \rho < R)
\end{aligned}\right\} \quad (22)\ .$$

These equations completely express the boundary conditions at $z = 0$, since in the present problem E_ρ and H_ϕ are the only transverse components of the $\underline{E}$ and $\underline{H}$ fields.

8 CURRENT AND VOLTAGE IN THE PRINCIPAL MODE

The next stage in the development is to obtain expressions for the "current" and "voltage" at the transverse plane $z = 0$. In the present case, since we are dealing with a two-conductor system, natural meanings can be attached to these quantities because the transverse field distribution arising from a TEM mode is essentially independent of frequency. If we make use of this fact to reduce the problem to an electrostatic one, eqn (3a) becomes

curl $\underline{E} = \underline{0}$ which implies that $\underline{E} = -\,\mathrm{grad}\ V$. From this we obtain

$$\text{voltage} = \int_{\text{inner radius}}^{\text{outer radius}} \left[\mathcal{E}_\rho\right] d\rho \quad (23a)$$

taking the inner conductor as positive with respect to the outer one, where [ ] means "the TEM part of".

In electrostatic terms, this corresponds to making the inner conductor the line and the outer one the return. The current in the inner is then positive, and is related to the magnetic field it produces by the most general electrostatic form of eqn (3c), which is (see [2])

curl $\underline{H}$ = $\underline{J}$. Applying Stokes' theorem (see [2]) to this, taking Φ as a simple closed curve around the inner conductor lying in the plane $z = 0$ and taking S as the portion of $z = 0$ bounded by Φ, we have

$$\text{current} = \iint_S \underline{J}\cdot\underline{dS} = \oint_\Phi \left[\underline{H}_{(z=0)}\right]\cdot\underline{dl} = \oint_\Phi \left[\mathcal{H}_\phi\right]\rho d\phi \qquad (23b)$$

in terms of the elementary vector area $\underline{dS}$ of S and the elementary vector length $\underline{dl}$ of Φ. The use of $\mathcal{H}_\phi$ does not mean that this current is flowing in the $z = 0$ plane; it is the z-directed current flowing into or out of that plane (in the positive z-direction).

These definitions of voltage and current can readily be generalised to other two-conductor systems. For one-conductor systems, however, there are no "natural" definitions available, and it is necessary to appeal to arbitrary definitions based on reflection coefficients; the fundamental relation here is that giving the reflection coefficient Γ of an admittance Y normalised to a characteristic admittance Y_o (see [1], [3])

$$\frac{Y}{Y_o} = \frac{1 - \Gamma}{1 + \Gamma} \quad .$$

This relation can be used in conjunction with eqns (21) and (22) to obtain Y, by the method described in [1]; but this method involves several complications, in particular the introduction of the concepts of reaction and complex power, and the implied assumption that the mode constants α_i, β_j are all real (or at any rate can be made so by multiplying them all by the same complex number). We will avoid these problems by continuing the development independently of [1] (an example of the connection between the two treatments will be given later).

9 *THE INTEGRAL RELATIONS*

At this point we introduce four essential integral relations between the mode functions. It is well known, and is proved in [1], [2] and [3], that the

mode functions appearing in such equations as (17) are orthogonal with respect to the cross-sectional area of the generalised cylindrical system to which they relate; that is, if $f_1(\rho,\phi)$ and $f_2(\rho,\phi)$ are two such transverse mode functions, then

$$\iint_{\text{cross-section}}^{\text{transverse}} f_1(\rho,\phi)f_2(\rho,\phi)dA \; = \; 0 \quad \text{unless } f_1 = f_2 \quad .$$

In the present case, where dA is $\rho d\rho d\phi$ and the functions are all independent of ϕ, this suggests that integrals of the form

$$\int f_1(\rho)f_2(\rho)\rho d\rho$$

will be important. Regarding the TEM mode function as being in a class of its own, we obtain four such fundamental integrals, all of which can be put in closed form using [4] (similar results also exist in other problems with generalised cylinders, but they cannot usually be expressed so conveniently; see [2]). The relations are

$$\int\left(\frac{1}{\rho}\right)^2 \rho d\rho \; = \; \ln \rho \tag{24a}$$

$$\int Z_1^2(\theta\rho)\rho d\rho \; = \; \frac{1}{2}\rho^2(Z_1^2(\theta\rho) - Z_o(\theta\rho)Z_2(\theta\rho)) \tag{24b}$$

$$\int\left(\frac{1}{\rho}\right)Z_1(\theta\rho)\rho d\rho \; = \; -\frac{1}{\theta}Z_o(\theta\rho) \tag{24c}$$

$$\underset{(\theta\neq\phi)}{\int Z_1(\theta\rho)Z_1(\phi\rho)\rho d\rho} \; = \; \frac{1}{(\theta^2 - \phi^2)}(\phi\rho Z_1(\theta\rho)Z_o(\phi\rho) - \theta\rho Z_1(\phi\rho)Z_o(\theta\rho)) \tag{24d}$$

Eqns (24c) and (24d) are the orthogonality relations properly so called, and when suitable limits are inserted and eqn (14) used it can be seen that the mode functions are indeed orthogonal. (The definitions of the Z functions are taken from eqns (13).) Eqn (24d) is valid even if θ and ϕ belong to different families (one a k_{Ai} and the other a k_{Bj}) although orthogonality does not then hold good.

From eqns (14), (21), (22), (23) and (24c) we can now determine the voltages across the discontinuity plane, and the currents flowing into it and out of it, on the A and B sides. In an obvious notation we have

$$V_A = \int_{r_A}^{R} \left[\mathcal{E}_\rho\right] d\rho = \int_{r_A}^{R} \frac{1}{\rho}\,\alpha_o(1 + \Gamma_A)d\rho = \alpha_o(1 + \Gamma_A)\ln(R/r_A) \qquad (25a)$$

and

$$V_B = \int_{r_B}^{R} \left[\mathcal{E}_\rho\right] d\rho = \int_{r_B}^{R} \frac{1}{\rho}\,\beta_o(1 + \Gamma_B)d\rho = \beta_o(1 + \Gamma_B)\ln(R/r_B) \quad . \qquad (25b)$$

Also

$$V_A = \int_{r_A}^{R} \frac{1}{\rho}\,\alpha_o(1 + \Gamma_A)d\rho = \int_{r_A}^{R} \left\{\frac{1}{\rho}\,\alpha_o(1 + \Gamma_A) + \sum_{i=1}^{\infty} \alpha_i Z_1(k_{Ai}\rho)\right\} d\rho$$

(using eqns (24c) and (14), and inverting the order of integration and summation as usual in Fourier analysis)

$$= \int_{r_A}^{R} \mathcal{E}_\rho d\rho = \int_{r_B}^{R} \mathcal{E}_\rho d\rho = \int_{r_B}^{R} \left\{\frac{1}{\rho}\,\beta_o(1 + \Gamma_B) + \sum_{j=1}^{\infty} \beta_j Z_1(k_{Bj}\rho)\right\} d\rho$$

from eqn (21)

$$= \int_{r_B}^{R} \frac{1}{\rho}\,\beta_o(1 + \Gamma_B)d\rho \qquad \text{from eqns (24c) and (14)}$$

$$= V_B = V \ \text{(say)} \quad . \qquad (25c)$$

Again

$$I_A = \int_0^{2\pi} \left[\mathcal{H}_\phi\right] \rho d\phi = \int_0^{2\pi} \frac{1}{\eta_A \rho} \alpha_o (1 - \Gamma_A) \rho d\phi = \frac{2\pi}{\eta_A} \alpha_o (1 - \Gamma_A) \tag{26a}$$

$$I_B = \int_0^{2\pi} \left[\mathcal{H}_\phi\right] \rho d\phi = \int_0^{2\pi} \frac{1}{\eta_B \rho} \beta_o (-1 + \Gamma_B) \rho d\phi = -\frac{2\pi}{\eta_B} \beta_o (1 - \Gamma_B) \quad . \tag{26b}$$

Since $V_A = V_B$, it is clear that the equivalent circuit representing the discontinuity must consist simply of a shunt admittance; any other components would introduce a voltage discontinuity across $z = 0$. (In the truncated system (Fig 1a) V_B has no meaning; but in this case there is no free propagation beyond $z = 0$ and the equivalent circuit of the whole system is terminated there, so we again arrive at a shunt admittance as the representation of the discontinuity.) A current I_A arrives at $z = 0$ from the A side and a current I_B leaves $z = 0$ from the B side (since I_A and I_B are both flowing in the positive z-direction, by eqn (23b)). The difference of these two currents must be considered to be driven through the shunt admittance, Y, by the voltage $V_A = V_B = V$; so

$$YV = I_A - I_B \quad . \tag{27}$$

This equation also applies to the system of Fig 1a, since then $\beta_o = 0$ (as remarked after eqn (17a)) and so from eqn (26b) $I_B = 0$; eqn (27) then states that the entire current I_A arriving at $z = 0$ from the A-side is diverted through the shunt admittance Y, which is clearly true since in this system the equivalent circuit is a terminating impedance at $z = 0$. In this case the simplified form of eqn (27), $YV = I_A$, may be reduced using eqns (25) and (26) to

$$Y = \left(\frac{2\pi}{\eta_A \ln (R/r_A)}\right)\left(\frac{1 - \Gamma_A}{1 + \Gamma_A}\right) \quad ,$$

which illustrates the general relation from [1] stated previously,

$$\frac{Y}{Y_o} = \frac{1 - \Gamma}{1 + \Gamma} \quad ,$$

with the characteristic admittance Y_o assuming its transmission-line value.

If we substitute into eqn (27) from eqns (25) and (26), we obtain

$$Y\left(\int_{r_A}^{R} \mathcal{E}_{\rho'}d\rho'\right) = YV = \frac{2\pi}{\eta_A}\alpha_o(1 - \Gamma_A) + \frac{2\pi}{\eta_B}\beta_o(1 - \Gamma_B)$$

$$= 2\pi\rho\left\{\sum_{i=1}^{\infty}\left(\frac{j\omega\varepsilon_A}{\gamma_{Ai}}\right)\alpha_i Z_1(k_{Ai}\rho) + \sum_{j=1}^{\infty}\left(\frac{j\omega\varepsilon_B}{\gamma_{Bj}}\right)\beta_j Z_1(k_{Bj}\rho)\right\} \qquad (28)$$

using eqn (22) and introducing the dummy variable ρ' to avoid confusion with ρ (which may assume any value in the range $r_A < \rho < R$; for ρ values within this range the right-hand side of eqn (28) is actually a constant, although it looks like a function of ρ). The integrations in this equation and in the equations which follow extend over the gap between the conductors at $z = 0$. In the present case the dimensions of this gap coincide with those of one of the pairs of cylinders, but this need not be the case (for instance, there might be a diaphragm at $z = 0$; or the cylinders might be stepped so that their radii both increased, which would create a gap smaller than the spacing between the cylinders on either side, as in Fig 1e). The integrations extend formally over the area of the gap, and to apply the method of this memorandum in general it is necessary to replace the single integrals in eqns (25), (26) and (28) by double integrals; the formulation in this case is discussed in [1]. However, the rotational symmetry of the systems represented in Figs 1 allows us to deal with them using the simpler single integrals.

11 EXPRESSION FOR THE SHUNT ADMITTANCE

We now express the quantities α_i and β_j in terms of $\mathcal{E}_\rho$. From eqn (21) we have

$$\int_{r_A}^{R} \mathcal{E}_\rho Z_1(k_{Ai}\rho)\rho\, d\rho = \int_{r_A}^{R}\left\{\frac{1}{\rho}\alpha_o(1 + \Gamma_A) + \sum_{i'=1}^{\infty}\alpha_{i'}Z_1(k_{Ai'}\rho)\right\}Z_1(k_{Ai}\rho)\rho\, d\rho$$

(introducing the dummy suffix i' to avoid confusion with i)

$$= \int_{r_A}^{R}\alpha_o(1 + \Gamma_A)Z_1(k_{Ai}\rho)d\rho + \sum_{i'=1}^{\infty}\int_{r_A}^{R}\alpha_{i'}Z_1(k_{Ai'}\rho)Z_1(k_{Ai}\rho)\rho\, d\rho$$

(interchanging the order of integration and summation)

$$= \alpha_o(1 + \Gamma_A)\left[-\frac{1}{k_{Ai}} Z_o(k_{Ai}\rho)\right]_{r_A}^{R} +$$

$$+ \sum_{\substack{i'=1 \\ i'\neq i}}^{\infty} \alpha_{i'}\left[\frac{(k_{Ai}\rho Z_1(k_{Ai'}\rho) Z_o(k_{Ai}\rho) - k_{Ai'}\rho Z_1(k_{Ai}\rho) Z_o(k_{Ai'}\rho))}{(k_{Ai'}^2 - k_{Ai}^2)}\right]_{r_A}^{R} +$$

$$+ \alpha_i \int_{r_A}^{R} Z_1^2(k_{Ai}\rho)\rho d\rho$$

(using eqns (24c) and (24d))

$$= \alpha_i \int_{r_A}^{R} Z_1^2(k_{Ai}\rho)\rho d\rho \qquad (29a)$$

(using eqn (14)).

This is an example of the orthogonality property of the mode functions. The integral on the last line of eqn (29a) can also be written in closed form using eqn (24b), but we shall leave it as it is for the time being. We can also write, using eqn (21),

$$\int_{r_A}^{R} \mathcal{E}_\rho Z_1(k_{Bj}\rho)\rho d\rho = \int_{r_B}^{R} \mathcal{E}_\rho Z_1(k_{Bj}\rho)\rho d\rho \qquad (29b)$$

In terms of our earlier discussion of these integrations, the last equation transforms an implicit double integral over the gap at $z = 0$ to one over the whole cross-section there. Since the cross-section consists entirely of the gap and a flat section (or sections) of perfect conductor, such a transformation is always possible, even when the system is not rotationally-symmetrical (see [1]). In systems like that of Fig 1e, it is necessary to transform in this way on both sides of $z = 0$, and we then have to introduce a third set of orthogonal functions (orthogonal over the gap, instead of over one of the generalised

cylindrical cross-sections). In the present problem, fortunately, this complication is unnecessary.

From eqns (29b) and (21), we now obtain

$$\int_{r_A}^{R} \mathcal{E}_\rho Z_1(k_{Bj}\rho)\rho d\rho = \int_{r_B}^{R} \left\{ \frac{1}{\rho}\, \beta_o(1 + \Gamma_B) + \sum_{j'=1}^{\infty} \beta_{j'} Z_1(k_{Bj'}\rho) \right\} Z_1(k_{Bj}\rho)\rho d\rho$$

$$= \beta_j \int_{r_B}^{R} Z_1^2(k_{Bj}\rho)\rho d\rho \tag{29c}$$

on repeating the steps in the derivation of eqn (29a). Substituting eqns (29) into eqn (28) gives, using the dummy ρ' as before,

$$\frac{Y}{2\pi}\,\frac{1}{\rho}\int_{r_A}^{R} \mathcal{E}_{\rho'} d\rho' = \sum_{i=1}^{\infty}\left(\frac{j\omega\varepsilon_A}{\gamma_{Ai}}\right)\left(\frac{\int_{r_A}^{R} \mathcal{E}_{\rho'} Z_1(k_{Ai}\rho')\rho' d\rho'}{\int_{r_A}^{R} Z_1^2(k_{Ai}\rho')\rho' d\rho'}\right) Z_1(k_{Ai}\rho) \; +$$

$$+ \sum_{j=1}^{\infty}\left(\frac{j\omega\varepsilon_B}{\gamma_{Bj}}\right)\left(\frac{\int_{r_A}^{R} \mathcal{E}_{\rho'} Z_1(k_{Bj}\rho')\rho' d\rho'}{\int_{r_B}^{R} Z_1^2(k_{Bj}\rho')\rho' d\rho'}\right) Z_1(k_{Bj}\rho) \quad . \tag{30}$$

Finally, if we multiply this equation throughout by $\mathcal{E}_\rho \rho$ and integrate across the gap (which in our reduced terms is equivalent to multiplying by $\mathcal{E}_\rho$ and integrating over the area of the gap) we obtain, on interchanging the order of integration and summation,

$$\frac{Y}{2\pi}\left(\int_{r_A}^{R} \mathcal{E}_\rho d\rho\right)^2 = \sum_{i=1}^{\infty}\left(\frac{j\omega\varepsilon_A}{\gamma_{Ai}}\right)\frac{\left(\int_{r_A}^{R} \mathcal{E}_\rho Z_1(k_{Ai}\rho)\rho d\rho\right)^2}{\left(\int_{r_A}^{R} Z_1^2(k_{Ai}\rho)\rho d\rho\right)} \; +$$

$$+\sum_{j=1}^{\infty}\left(\frac{j\omega\epsilon_B}{\gamma_{Bj}}\right)\frac{\left(\int_{r_A}^{R}\mathcal{E}_\rho Z_1(k_{Bj}\rho)\rho d\rho\right)^2}{\left(\int_{r_B}^{R}Z_1^2(k_{Bj}\rho)\rho d\rho\right)} \qquad . \tag{31}$$

12 STATIONARITY AND OTHER VARIATIONAL PROPERTIES

Eqn (31) is a special case of a general formula derived in [1], a so-called variational formula. The properties of variational formulae are discussed in [1] and [2], where it is shown that they have the remarkable property of stationarity. In our case this means that, if the true field $\mathcal{E}_\rho$ which appears in eqn (30) is replaced by an arbitrary function $\hat{\mathcal{E}}_\rho$, then, provided the arbitrary function (like the true field) satisfies the proper boundary conditions at the edges of the gap,

$$\frac{\partial Y}{\partial \hat{\mathcal{E}}_\rho} = 0 \qquad \text{when } \hat{\mathcal{E}}_\rho = \mathcal{E}_\rho \qquad .$$

A proof of this property is most conveniently given using the development of the calculus of variations by Courant and Hilbert [5]. They show that a double integral of the form

$$\int_a^b\int_a^b K(\rho,\rho')\,\hat{\mathcal{E}}_\rho(\rho)\,\hat{\mathcal{E}}_\rho(\rho')d\rho d\rho' \qquad , \tag{32a}$$

where a and b are constants, and $K(\rho,\rho')$ is a fixed function satisfying conventional conditions of continuity and differentiability and possessing the symmetry property $K(\rho,\rho') = K(\rho',\rho)$, will be stationary with respect to changes in the functional form of $\hat{\mathcal{E}}_\rho$ if its functional gradient, which is

$$2\int_a^b K(\rho,\rho')\,\hat{\mathcal{E}}_\rho(\rho')d\rho' \qquad , \tag{32b}$$

vanishes for all values of ρ in the range $a < \rho < b$. The squared integrals in

eqn (31) can be written as double integrals of the form (32a); freely interchanging the orders of integration and summation, and replacing ρ by ρ' where appropriate, we find that eqn (31) and the equation obtained by setting expression (32a) equal to zero can be identified with each other if we choose $a = r_A$, $b = R$ and

$$K(\rho,\rho') = \sum_{i=1}^{\infty}\left(\frac{j\omega\varepsilon_A}{\gamma_{Ai}}\right)\left(\int_{r_A}^{R} Z_1^2(k_{Ai}\rho)\rho d\rho\right)^{-1} Z_1(k_{Ai}\rho)Z_1(k_{Ai}\rho')\rho\rho' +$$

$$+ \sum_{j=1}^{\infty}\left(\frac{j\omega\varepsilon_B}{\gamma_{Bj}}\right)\left(\int_{r_B}^{R} Z_1^2(k_{Bj}\rho)\rho d\rho\right)^{-1} Z_1(k_{Bj}\rho)Z_1(k_{Bj}\rho')\rho\rho' -$$

$$- \frac{Y}{2\pi} \tag{32c}$$

which has the required symmetry property under interchange of ρ and ρ'. It then turns out that the functional gradient, expression (32b), vanishes when $\hat{\mathcal{E}}_\rho = \mathcal{E}_\rho$, because of eqn (30).

Now, suppose we try to find an approximate value of Y by using an approximate function $\hat{\mathcal{E}}_\rho$ to represent $\mathcal{E}_\rho$ in eqn (31). If the error in the approximation $\hat{\mathcal{E}}_\rho$ is of the first order of small quantities, the error in the computed approximate Y will only be of the second order of smallness, because of the vanishing partial derivative with respect to $\hat{\mathcal{E}}_\rho$ at $\hat{\mathcal{E}}_\rho = \mathcal{E}_\rho$. This is an important consequence of stationarity; but still more important is the fact that stationarity gives us a simple method of constructing good approximations $\hat{\mathcal{E}}_\rho$. If we imagine generalised cylinders to be obtained by translating along the z-axis the boundary curves of the gap in the plane z = 0, then the mode functions appropriate to these cylinders satisfy the proper boundary conditions in the plane of the gap, and so any linear combination of these mode functions is a possible approximation $\hat{\mathcal{E}}_\rho$. If we choose a finite linear combination, use it as an approximation $\hat{\mathcal{E}}_\rho$ to represent $\mathcal{E}_\rho$ in eqn (31), and adjust all the linear multipliers to satisfy the stationarity condition $(\partial Y/\partial\hat{\mathcal{E}}_\rho) = 0$, we shall obtain a second-order approximation to the value of Y. By using more and more mode functions in linear combination, we can obtain a sequence of approximations to the value of Y. Since the mode functions form a mathematically

"complete" set (see [2] or [5]), we can represent the true $\mathcal{E}_\rho$ as accurately as we please (at least in the mean, which is all that matters) by taking a sufficient number of them in a suitable linear combination; so, if the sequence of approximations tends to a limit (which we shall show to be the case), the limit will be the true value of Y. This is the well-known Ritz procedure (discussed in [1], [2] and [5]).

13 APPROXIMATIONS BY THE RITZ PROCEDURE

We now apply the procedure just described. Let $\mathcal{E}_\rho$ in eqn (31) be replaced by an approximation $\hat{\mathcal{E}}_\rho$ of the form

$$\hat{\mathcal{E}}_\rho = \frac{1}{\rho}\hat{\alpha}_o + \sum_{i'=1}^{N} k_{Ai'}\hat{\alpha}_{i'} Z_1(k_{Ai'}\rho) \tag{33}$$

since the generalised cylinders constructed by translating along the z-axis the gap in the plane z = 0 are, in this case, just the same as the cylinders on the A-side (if they were different from the cylinders on the A and B sides, as in the system of Fig 1e, it would be necessary to introduce a third set of orthogonal mode functions, but in the present problem this complication does not arise). The factor $k_{Ai'}$ in each term of the sum in eqn (33) is inserted for future convenience, and N is an unspecified positive integer. We then have, from eqns (24) and (14),

$$\int_{r_A}^{R} Z_1^2(k_{Ai}\rho)\rho d\rho = \frac{1}{2}\left(R^2 Z_1^2(k_{Ai}R) - r_A^2 Z_1^2(k_{Ai}r_A)\right) \tag{34a}$$

$$\int_{r_B}^{R} Z_1^2(k_{Bj}\rho)\rho d\rho = \frac{1}{2}\left(R^2 Z_1^2(k_{Bj}R) - r_B^2 Z_1^2(k_{Bj}r_B)\right) \tag{34b}$$

$$\int_{r_A}^{R} \hat{\mathcal{E}}_\rho d\rho = \hat{\alpha}_o \ln (R/r_A) \tag{34c}$$

$$\int_{r_A}^{R} \hat{\mathcal{E}}_\rho Z_1(k_{Ai}\rho)\rho d\rho = \frac{1}{2} k_{Ai}\hat{\alpha}_i (R^2 Z_1^2(k_{Ai}R) - r_A^2 Z_1^2(k_{Ai}r_A)) \quad (i \leq N)$$

$$= 0 \quad (i > N) \qquad (34d)$$

$$\int_{r_A}^{R} \hat{\mathcal{E}}_\rho Z_1(k_{Bj}\rho)\rho d\rho = \hat{\alpha}_o \frac{Z_o(k_{Bj}r_A)}{k_{Bj}} + \sum_{i'=1}^{N} \left(\frac{k_{Ai'}\hat{\alpha}_{i'}}{k_{Bj}^2 - k_{Ai'}^2}\right) k_{Bj} r_A Z_1(k_{Ai'}r_A) Z_o(k_{Bj}r_A) \; . \qquad (34e)$$

On making use of eqns (33) and (34) in eqn (31), we obtain an approximation to Y, $\hat{Y}$, given by

$$\frac{\hat{Y}}{2\pi j\omega}(\hat{\alpha}_o \ln (R/r_A))^2 = \sum_{i=1}^{N} \left(\frac{\varepsilon_A}{2\gamma_{Ai}}\right) \left(R^2 Z_1^2(k_{Ai}R) - r_A^2 Z_1^2(k_{Ai}r_A)\right) k_{Ai}^2 \hat{\alpha}_i^2 +$$

$$+ \sum_{j=1}^{\infty} \left(\frac{\varepsilon_B}{\gamma_{Bj}}\right) \left(\frac{2}{R^2 Z_1^2(k_{Bj}R) - r_B^2 Z_1^2(k_{Bj}r_B)}\right) \left(\frac{Z_o(k_{Bj}r_A)}{k_{Bj}}\right)^2 \times$$

$$\times \left\{ \hat{\alpha}_o + \sum_{i'=1}^{N} \left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai'}^2}\right) k_{Ai'} r_A Z_1(k_{Ai'}r_A) \hat{\alpha}_{i'} \right\}^2$$

which becomes, on interchanging the order of summation and making some convenient changes in the dummy variables,

$$\frac{\hat{Y}}{2\pi j\omega}(\hat{\alpha}_o \ln (R/r_A))^2 = \hat{\alpha}_o^2 \sum_{j=1}^{\infty} \left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right) \left(\frac{4 Z_o^2(k_{Bj}r_A)}{(k_{Bj}RZ_1(k_{Bj}R))^2 - (k_{Bj}r_B Z_1(k_{Bj}r_B))^2}\right) +$$

$$+ \sum_{i=1}^{N} \hat{\alpha}_i^2 \left[\frac{\varepsilon_A}{2\gamma_{Ai}} \left((k_{Ai}RZ_1(k_{Ai}R))^2 - (k_{Ai}r_A Z_1(k_{Ai}r_A))^2\right)\right] +$$

$$+ 2\hat{\alpha}_o \sum_{i=1}^{N} \hat{\alpha}_i \left[(k_{Ai} r_A Z_1(k_{Ai} r_A)) \sum_{j=1}^{\infty} \left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2} \right) \left(\frac{\varepsilon_B}{2\gamma_{Bj}} \right) \times \right.$$

$$\left. \times \left(\frac{4Z_o^2(k_{Bj} r_A)}{(k_{Bj} R Z_1(k_{Bj} R))^2 - (k_{Bj} r_B Z_1(k_{Bj} r_B))^2} \right) \right] +$$

$$+ \sum_{i=1}^{N} \sum_{i'=1}^{N} \hat{\alpha}_i \hat{\alpha}_{i'} \left[(k_{Ai} r_A Z_1(k_{Ai} r_A))(k_{Ai'} r_A Z_1(k_{Ai'} r_A)) \sum_{j=1}^{\infty} \left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2} \right) \left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai'}^2} \right) \right.$$

$$\left. \times \left(\frac{\varepsilon_B}{2\gamma_{Bj}} \right) \left(\frac{4Z_o^2(k_{Bj} r_A)}{(k_{Bj} R Z_1(k_{Bj} R))^2 - (k_{Bj} r_B Z_1(k_{Bj} r_B))^2} \right) \right] \quad (35)$$

14 SIMPLIFICATION OF THE APPROXIMATE EXPRESSION

Eqn (35) admits a number of simplifications. In the first place, it is clear that only the ratios of the $\hat{\alpha}$'s are significant, since the equation is homogeneous in them. It is convenient to write

$$\hat{\alpha}_i (k_{Ai} r_A Z_1(k_{Ai} r_A)) = - a_i \hat{\alpha}_o \quad (36a)$$

where the a_i are the new unknown multipliers.
Secondly, if ξ denotes one of the quantities $(k_{Ai} r_A)$, $(k_{Bj} r_B)$, $(k_{Ai} R)$, $(k_{Bj} R)$, we have from eqns (12) and (13)

$$Z_1(\xi) = J_1(\xi) - \frac{J_o(\xi)}{Y_o(\xi)} Y_1(\xi)$$

But for any complex number w, it is true that

$$J_1(w)Y_o(w) - J_o(w)Y_1(w) = J_o(w)Y_o'(w) - Y_o(w)J_o'(w) = \frac{2}{\pi w}$$

(after [2] or [4])
provided $w \neq 0$, where the prime denotes differentiation; so

$$Z_1(\xi) = 2/(\pi\xi Y_o(\xi)) \tag{36b}$$

($\xi \neq 0$, because the k_{Ai}, the k_{Bj} and r_A, r_B and R are all strictly positive).

We therefore have, after dividing by $\hat{\alpha}_o^2$,

$$\frac{\hat{Y}}{2\pi j\omega}\,(\ln\,(R/r_A))^2 = \sum_{j=1}^{\infty}\left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right)\left(\frac{4Z_o^2(k_{Bj}r_A)}{(k_{Bj}RZ_1(k_{Bj}R))^2 - (k_{Bj}r_BZ_1(k_{Bj}r_B))^2}\right) +$$

$$+\sum_{i=1}^{N} a_i^2\left[\frac{\varepsilon_A}{2\gamma_{Ai}}\left(\left(\frac{k_{Ai}RZ_1(k_{Ai}R)}{k_{Ai}r_AZ_1(k_{Ai}r_A)}\right)^2 - 1\right)\right] -$$

$$-\,2\sum_{i=1}^{N} a_i\left[\sum_{j=1}^{\infty}\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2}\right)\left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right)\left(\frac{4Z_o^2(k_{Bj}r_A)}{(k_{Bj}RZ_1(k_{Bj}R))^2 - (k_{Bj}r_BZ_1(k_{Bj}r_B))^2}\right)\right] +$$

$$+\sum_{i=1}^{N}\sum_{i'=1}^{N} a_ia_{i'}\left[\sum_{j=1}^{\infty}\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2}\right)\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai'}^2}\right)\left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right)\times\right.$$

$$\left.\times\left(\frac{4Z_o^2(k_{Bj}r_A)}{(k_{Bj}RZ_1(k_{Bj}R))^2 - (k_{Bj}r_BZ_1(k_{Bj}r_B))^2}\right)\right]$$

(from eqn (36a))

$$= \sum_{j=1}^{\infty}\left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right)\left(\frac{\pi^2Z_o^2(k_{Bj}r_A)}{(Y_o(k_{Bj}R))^{-2} - (Y_o(k_{Bj}r_B))^{-2}}\right) + \sum_{i=1}^{N} a_i^2\left[\frac{\varepsilon_A}{2\gamma_{Ai}}\left(\left(\frac{Y_o(k_{Ai}r_A)}{Y_o(k_{Ai}R)}\right)^2 - 1\right)\right] -$$

$$-\,2\sum_{i=1}^{N} a_i\left[\sum_{j=1}^{\infty}\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2}\right)\left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right)\left(\frac{\pi^2Z_o^2(k_{Bj}r_A)}{(Y_o(k_{Bj}R))^{-2} - (Y_o(k_{Bj}r_B))^{-2}}\right)\right] +$$

$$+\sum_{i=1}^{N}\sum_{i'=1}^{N} a_ia_{i'}\left[\sum_{j=1}^{\infty}\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2}\right)\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai'}^2}\right)\left(\frac{\varepsilon_B}{2\gamma_{Bj}}\right)\times\right.$$

$$\left.\times\left(\frac{\pi^2Z_o^2(k_{Bj}r_A)}{(Y_o(k_{Bj}R))^{-2} - (Y_o(k_{Bj}r_B))^{-2}}\right)\right] \tag{37}$$

(from eqn (36b)) .

Again, from eqns (12) and (13) we have

$$Z_o(k_{Bj}r_A) = J_o(k_{Bj}r_A) - \frac{J_o(k_{Bj}R)}{Y_o(k_{Bj}R)} Y_o(k_{Bj}r_A)$$

so if we define

$$x_j = \frac{\left(J_o(k_{Bj}r_A)Y_o(k_{Bj}R) - J_o(k_{Bj}R)Y_o(k_{Bj}r_A)\right)^2}{\gamma_{Bj}\left(1 - \left(\frac{Y_o(k_{Bj}R)}{Y_o(k_{Bj}r_B)}\right)^2\right)} \tag{38a}$$

then

$$x_j = \frac{Z_o^2(k_{Bj}r_A)}{\gamma_{Bj}\left((Y_o(k_{Bj}R))^{-2} - (Y_o(k_{Bj}r_B))^{-2}\right)} \quad . \tag{38b}$$

If we also introduce the relative permeabilities and permittivities $\overline{\mu_A}$, $\overline{\mu_B}$, $\overline{\varepsilon_A}$, $\overline{\varepsilon_B}$, the absolute permeability and permittivity of free space μ_o, ε_o, the velocity of light in free space c and the well-known relation ([1], [2])

$$\mu_o \varepsilon_o c^2 = 1$$

we obtain, using eqn (6a),

$$k_A^2 = \omega^2 \, \overline{\mu_A} \, \overline{\varepsilon_A} \, / \, c^2 \tag{38c}$$

with a similar relation for k_B^2, and

$$\varepsilon_B = \overline{\varepsilon_B} \, / \, (\mu_o c^2) \quad . \tag{38d}$$

Introducing eqns (38b) and (38d) into eqn (37), along with the further definitions

$$q = \sum_{j=1}^{\infty} x_j \tag{38e}$$

$$s_i = \left(\frac{1}{\pi^2 \gamma_{Ai}}\right)\left(\frac{\overline{\varepsilon_A}}{\overline{\varepsilon_B}}\right)\left(\left(\frac{Y_o(k_{Ai} r_A)}{Y_o(k_{Ai} R)}\right)^2 - 1\right) \tag{38f}$$

$$t_i = \sum_{j=1}^{\infty}\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2}\right) x_j \tag{38g}$$

$$u_{i,i'} = u_{i',i} = \sum_{j=1}^{\infty}\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai}^2}\right)\left(\frac{k_{Bj}^2}{k_{Bj}^2 - k_{Ai'}^2}\right) x_j \quad , \tag{38h}$$

gives us

$$\frac{\hat{Y}}{j\omega} = \frac{\pi^3 \,\overline{\varepsilon_B}}{\mu_o c^2 (\ln\,(R/r_A))^2}\left\{q + \sum_{i=1}^{N} a_i^2 s_i - 2\sum_{i=1}^{N} a_i t_i + \sum_{i=1}^{N}\sum_{i'=1}^{N} a_i a_{i'} u_{i,i'}\right\} . \tag{39}$$

15 STATIONARY VALUE OF THE APPROXIMATE ADMITTANCE

Applying the Ritz procedure to determine the stationary value of $\hat{Y}$ in eqn (39) is now just a matter of taking its partial derivatives with respect to each of the a_i in turn and setting all the resulting expressions equal to zero. Taking note of the fact that $u_{i,i'} = u_{i',i}$ (eqn (38h)) it is easy to see that

$$\frac{\partial}{\partial a_\ell}\left\{q + \sum_{i=1}^{N} a_i^2 s_i - 2\sum_{i=1}^{N} a_i t_i + \sum_{i=1}^{N}\sum_{i'=1}^{N} a_i a_{i'} u_{i,i'}\right\}$$

$$= 2a_\ell s_\ell - 2t_\ell + 2a_\ell u_{\ell,\ell} + 2\sum_{\substack{i=1\\ i\neq\ell}}^{N} a_i u_{\ell,i}$$

and consequently the values of the a_i which make $\hat{Y}$ stationary are given by solving the simultaneous linear equations

$$(s_\ell + u_{\ell,\ell})\, {}^N a_\ell + \sum_{\substack{i=1\\ i\neq\ell}}^{N} u_{\ell,i}\, {}^N a_i = t_\ell \qquad (\ell = 1, 2, \ldots, N) \tag{40}$$

where a superscript N has been introduced on the a_i to stress that their best values depend on N. Substituting eqns (40) into eqn (39), and writing $\hat{Y}_N$ for the stationary value of $\hat{Y}$ (which depends on N, because the best values of the a_i do so) we obtain

$$\frac{\hat{Y}_N}{j\omega} = \frac{\pi^3 \overline{\varepsilon_B}}{\mu_o c^2 (\ln (R/r_A))^2} \left\{ q - \sum_{i=1}^{N} {}^N a_i t_i \right\} \tag{41}$$

because multiplying each of eqns (40) by the corresponding ${}^N a_\ell$ and then adding all the results together gives

$$\sum_{\ell=1}^{N} {}^N a_\ell^2 s_\ell + \sum_{\ell=1}^{N} \sum_{i=1}^{N} {}^N a_\ell \, {}^N a_i \, u_{\ell,i} = \sum_{\ell=1}^{N} {}^N a_\ell t_\ell \quad .$$

It can be seen that the solutions of eqns (40) are real numbers (not complex ones), because the definitions in eqns (38) show that the known quantities in eqns (40) are all real. It is then clear, from eqns (36a) and (33), that the approximate field $\hat{\mathcal{E}}_\rho$ is real, apart possibly from the common factor $\hat{\alpha}_o$ (this implies that the fields due to the individual modes are all in phase with each other, which is a quite reasonable physical condition for stationarity). Since $\hat{\mathcal{E}}_\rho$ can approach $\mathcal{E}_\rho$ as closely as we please, and since the factor $\hat{\alpha}_o$ is of no significance in any expression which is homogeneous in $\hat{\mathcal{E}}_\rho$, it then follows that $\mathcal{E}_\rho$ in eqn (31) can be taken as a pure real, and consequently that $Y/(j\omega)$ is a pure real. Moreover, this pure real is positive, because when eqn (31) is rewritten as an expression for $Y/(j\omega)$ it becomes apparent that every quantity in the expression is a perfect square or otherwise known to be positive (including the integrals in the denominators, which have non-negative integrands). It is clearly also true, by the same argument, that every $\hat{Y}_N/(j\omega)$ is a positive real. It follows, therefore, that the quadratic form in the a_i for $\hat{Y}$, eqn (39), must be positive-definite, and that it attains an absolute minimum when the best values given by eqn (40) are inserted in it. If N is increased to (N + 1), then in terms of the new approximate $\hat{\mathcal{E}}_\rho$ the old stationary field is no longer stationary, and when it is made so by replacing the ${}^N a_i$ by the ${}^{(N+1)} a_i$ it attains a new absolute minimum which must be lower than the old one. Accordingly the $\hat{Y}_N/(j\omega)$ form a monotonic decreasing real sequence in N, which is bounded below (since every member of it is positive); and so, since

$\hat{\mathcal{E}}_\rho$ can be made to approach $\mathcal{E}_\rho$ as closely as we please by taking N large enough, it follows that, as N tends to infinity, $\hat{Y}_N/(j\omega)$ tends from above to the value given by eqn (31) when $\hat{\mathcal{E}}_\rho = \mathcal{E}_\rho$, which is $Y/(j\omega)$.

16 THE DISCONTINUITY CAPACITANCE

We can now complete the formal processes by providing an interpretation of Y. We have already shown that $Y/(j\omega)$ is a positive real, and since the admittance of a capacitance is $(j\omega C)$ when the time variation of the fields is as $\exp(j\omega t)$ (as assumed at the beginning) it is convenient to interpret Y as the admittance of a capacitance $Y/(j\omega)$. This becomes not only convenient but also natural when we note, from eqns (10), (15c), (38), (39) and (40), that $Y/(j\omega)$ depends only slightly on the frequency (through the influence of k_A^2 and k_B^2 on the γ_{Ai} and γ_{Bj}) when the frequency is well below the upper critical frequency defined earlier, and that $Y/(j\omega)$ attains a definite (positive) value when the frequency is zero (from eqns (15c) and (38c) we then have simply $\underline{\gamma_{Ai}} = \underline{k_{Ai}}$, $\gamma_{Bj} = k_{Bj}$ and there are no numerical difficulties). Also, multiplying $\overline{\varepsilon_A}$ and $\overline{\varepsilon_B}$ simultaneously by a given factor causes the zero-frequency value of $Y/(j\omega)$ to increase by the same given factor, a result which remains approximately true at non-zero frequencies well below the upper critical frequency. These are all characteristics of a slightly imperfect capacitance; so we conclude that the equivalent circuit of a system like that of Fig 1a or Fig 1b should be a shunt capacitance, having a value given by $\lim_{N\to\infty} (\hat{Y}_N/(j\omega))$ with $\hat{Y}_N$ defined by eqn (41). A similar conclusion holds good for the other systems of Figs 1.

Most of the theory of the last few sections can be generalised to more complicated systems, as discussed in [1]; in particular it is nearly always possible to establish a variational equation like eqn (31), although the generalised variational equation will not necessarily lead to a positive-definite quadratic form. If no such form can be found, the smoothness and speed of convergence to the true value will be impaired, but convergence will still occur.

The special case represented by Fig 1a can be treated by allowing r_B to tend to 0 in the formulae derived above; eqn (10b) reduces to eqn (20a), and the functions $Z_{0/1/2}(k_{Bj}\rho)$ reduce to $J_{0/1/2}(k_{Bj}\rho)$, from eqns (12) and (13) and the discussion preceding them (note that $Y_o(k_{Bj}r_B)$ tends to infinity as r_B tends to 0). Eqn (38a) then assumes the special form

$$x_j = \frac{(J_o(k_{Bj}r_A)Y_o(k_{Bj}R))^2}{\gamma_{Bj}} \tag{42}$$

No other changes are necessary.

17 PRACTICAL DETAILS OF THE CALCULATION

It only remains to say how eqns (10), (15c), (20a), (38), (40), (41) and (42) should be used in practice to compute the shunt discontinuity capacitance.

The infinite sums in eqns (38) are naturally truncated to large finite upper limits. If the last term retained is the M^{th}, the error committed by truncation is of the order of M^{-2}, as can be shown by the following argument. Computing some of the sums for typical cases shows that the most important term in the brace brackets in eqn (41) is the leading term q, as we expect from the fact that when N = 0 the expression in the brace brackets actually reduces to q. It follows from eqn (38e) that we can estimate the truncation error by simply considering eqns (38a) and (42). The discussion in [2] shows that the large values of k_{Bj} are of the asymptotic form $j\pi/(R - r_B)$ (this formula actually holds fairly well also for small j, and was used to derive the approximate formulae for the critical frequencies). From [2] or [4], $J_o(z)$ and $Y_o(z)$ fall off like $z^{-\frac{1}{2}}$ for large z, and from eqn (15c) γ_{Bj} is asymptotic to k_{Bj} for large j. It follows from eqns (38a) and (42) that the x_j fall off like j^{-3} for large j, and the relative error committed by stopping at some large j equal to M is therefore of the order of

$$\left(\sum_{j=M+1}^{\infty} j^{-3}\right) \Bigg/ \left(\sum_{j=1}^{\infty} j^{-3}\right)$$

or of the order of M^{-2}, using integrals to approximate the sums. The error can therefore be reduced to something of the order of 10^{-6} by taking M = 1000, as was done by Somlo in the program referred to in his paper [6]; this is acceptably small, both in its own right and in comparison with Somlo's estimate of the accuracy of the whole calculation (which appears to depend mainly on the extrapolation to infinite N). We therefore cut off all the infinite sums after the first 1000 terms.

The Bessel functions can be evaluated by many standard methods and we shall not discuss them in detail. It is worth noting, however, that since the contribution of the j^{th} term to each sum falls off in importance like j^{-3}, it is acceptable to use methods which give progressively worse accuracy in $J_o(z)$ and $Y_o(z)$ as z increases (such as the fast and fairly simple method based on a Chebyshev series in z for small z and an asymptotic expansion in z^{-1} for large z, butted together).

The quantities k_{Ai} and k_{Bj} can be computed from eqns (10) and (20a), using standard methods of finding zeros. Since it is necessary to find every zero in a given range once and only once, it is desirable to use a method which takes an upper and a lower limit and finds a zero between them, using Bessel function theory to make sure that there is one and only one zero between the given limits. Suitable limits may be obtained from the asymptotic formula quoted above, and its degenerate form for the system of Fig 1a which is obtainable to sufficient accuracy by putting $r_B = 0$ (see [2] and [4]). We then find that the solutions of eqns (10b) and (20a) each satisfy the relation

$$\frac{\pi}{(R - r_B)}(j - 0.3) < k_{Bj} < \frac{\pi}{(R - r_B)}(j + 0.3) \tag{43}$$

where the (somewhat arbitrary) choice of $\pm$ 0.3 is wide enough to include one zero and narrow enough to include only one (and that one the j^{th}). (From [4], there are no multiple zeros.) In the same way we can set limits to k_{Ai} for use with eqn (10a).

The solution of eqns (40) for a given N can also be obtained by standard methods. Because of the positive-definiteness of eqn (39), eqns (40) have a positive-definite coefficient matrix. This is easily proved by writing $a_i = \overline{a_i}/\overline{a_{N+1}}$, when the positive-definite expression in the brace brackets in eqn (39) assumes the form

$$\overline{a_{N+1}}^{\,-2}\left(\overline{a_{N+1}}^{\,2}\, q + \sum_{i=1}^{N} \overline{a_i}^{\,2}\, s_i - 2\sum_{i=1}^{N} \overline{a_i}\, t_i\, \overline{a_{N+1}} + \sum_{i=1}^{N}\sum_{i'=1}^{N} \overline{a_i}\,\overline{a_{i'}}\, u_{i,i'}\right) . \tag{44a}$$

The homogeneous quadratic form in brackets here is necessarily also positive-definite, and its associated matrix B, of order (N + 1) x (N + 1), is therefore positive-definite too and has elements $b_{i,i'}$ satisfying the conditions $b_{i,i'} = b_{i',i}$ and (from expression (44a))

$$\sum_{i=1}^{N+1}\sum_{i'=1}^{N+1} \overline{a_i}\, b_{i,i'}\, \overline{a_{i'}} = \overline{a_{N+1}}^{\,2}\, q + \sum_{i=1}^{N} \overline{a_i}^{\,2}\, s_i - 2\sum_{i=1}^{N} \overline{a_i}\, t_i\, \overline{a_{N+1}} + \sum_{i=1}^{N}\sum_{i'=1}^{N} \overline{a_i}\, u_{i,i'}\, \overline{a_{i'}} \quad . \tag{44b}$$

Equating corresponding terms and taking account of the fact that $u_{i,i'} = u_{i',i}$, we have

$$\left.\begin{array}{lll}
b_{i,i'} = b_{i',i} = u_{i,i'} = u_{i',i} & & (i \leqslant N,\ i' \leqslant N,\ i' \neq i) \\
b_{i,i} = s_i + u_{i,i} & & (i \leqslant N) \\
b_{N+1,i} = b_{i,N+1} = -\,t_i & & (i \leqslant N) \\
b_{N+1,N+1} = q \quad . & &
\end{array}\right\} \tag{44c}$$

From Sylvester's theorem (see Faddeev and Faddeeva [7]) all the leading sub-matrices of a positive-definite matrix are also positive-definite, in particular the N^{th} such sub-matrix of B, which consists of the elements $b_{i,i'}$ $(i \leqslant N,\ i' \leqslant N)$; and from eqns (44c), this sub-matrix is precisely the coefficient matrix of eqns (40). Special algorithms exist for solving a system of equations with a positive-definite coefficient matrix, which reduce by half the computational effort required; however, none of these algorithms were used in programming the theory developed above, for reasons which are explained later. The positive-definiteness has another and much more important consequence: it guarantees that eqns (40) always possess a unique solution (see [7]). This solution is not necessarily well-determined (eqns (38) show that if any one of the k_{Bj} is nearly equal to one of the k_{Ai}, some of the $u_{i,i'}$ and t_i will be very large and not accurately calculable), but in practice i is always small enough for this problem not to arise. The author's program includes

a check that no difference $(k_{Ai} - k_{Bj})$ is less than 10^{-3} of the k_{Ai} used in computing it, and this check is always satisfied in the numerical cases so far treated, which restricts the loss of significant-figure accuracy to an acceptable three figures.

With the information in this section it is a straight-forward matter to construct a computer program which will calculate $\hat{Y}_N/(j\omega)$ for any given N, using as data the values of ω, $\overline{\mu_A}$, $\overline{\mu_B}$, $\overline{\epsilon_A}$, $\overline{\epsilon_B}$, R, r_A, r_B, μ_o and c. The extrapolation to infinite N will now be discussed.

18 EXTRAPOLATION TO INFINITY

The final stage of the computation consists of determining the discontinuity capacitance C, $Y/(j\omega)$, from the values of $\hat{Y}_N/(j\omega)$ for a number of values of N. It is convenient and makes good mathematical sense to choose consecutive values of N; in [6], for instance, the values 28, 29, ..., 40 were used. The theory of [6] is derived from Whinnery et al [8]; the final equations arrived at by these authors are similar to eqns (38) to (41), but no reasons are given for any of the formal manipulations in [8], the variational property is not mentioned, and the results are rather difficult to use because they have not been simplified as fully as has been done here. (Some of the difficulties of use are mentioned in [6], and these difficulties and the inconveniently concise presentation in [8] made it desirable to develop the theory systematically, so as to throw light on the numerical behaviour of the approximate solutions). Because the program corresponding to [6] used the theory of [8], which is similar to that presented above, it was expected that the convergence would be similar, so an upper limit to the value of N was chosen at 40. This proved to be satisfactory, which was fortunate because the time and storage demands of the program would have made it impossible to use a substantially higher limit.

In [6], and in Bianco et al [9], extrapolation by a hyperbolic fit (plotting $\hat{Y}_N$ against N^{-1}) was described, but the program corresponding to [6] (obtained from the author of that paper) actually used an implementation (no 215 of the ACM Collected Algorithms [12])of the e_2 transformation described by Shanks [10], and this more powerful method was employed in the present work. It is most easily applied to sequences containing (4n + 1) consecutive elements, n being an integer; the number of elements in the set 28, 29, ..., 40 is of

this form with $n = 3$, which suggests that Somlo changed his procedure while he was writing his paper. (He might have been troubled by an error in the printing of algorithm 215, which at one point has an illegal symbol in place of a semi-colon). In any case, we are provided with a suitable value of n. This is not a trivial matter, because the reduction of $(4n + 1)$ elements to a single final value actually requires n applications of the e_2 transformation, and repeated application of any smoothing algorithm can induce instability when the differences between successive smoothed values become comparable with the errors in the values themselves; examination of the intermediate values produced in the author's program by the smoothing algorithm showed, in fact, that taking n greater than 3 would have caused this problem to show itself.

Values of C calculated using $N = 28, 29, \ldots\ldots, 40$ were found to agree with the tabulated values in [6] to a few parts in 10^4. The author's values were consistently higher, which suggests, in view of the fact that the true value of C is approached from above as N increases, that the values in [6] are more accurate. However, as the calculations actually performed by the two programs are similar, it seems that the estimated accuracy of a few parts in 10^5 for the values in [6] may not be attained. For practical purposes the discrepancy is too small to matter. It is also worth recording that agreement with the values in [6] to a few parts in 10^3 was achieved using only sets of $\hat{Y}_N/(j\omega)$ calculated using $N = 1, 2, \ldots\ldots, 5$ with one application of the e_2 transformation; this remarkably high accuracy is a manifestation of the fact that the $\hat{Y}_N$ are second-order approximations to the true value, Y.

19 OTHER COMPUTATIONAL POINTS

It can be seen from eqns (40) that the coefficient matrix in the linear equations appropriate to any given N contains, as a leading sub-matrix of itself, the coefficient matrix appropriate to any smaller value of N. One property of a system of equations having a positive-definite coefficient matrix is that it can be solved accurately (by Gaussian elimination or by resolving the coefficient matrix into triangular factors) without altering the order in which the equations are presented (see Wilkinson [11], and also [7]) so that solving eqns (40) for the largest relevant N does most of the work of solving them for all smaller values of N; if the solution is undertaken by triangular factorisation, for instance, the triangular factors of the N'^{th} leading sub-matrix of the matrix for the largest N are simply the N'^{th} leading

sub-matrices of the triangular factors of the matrix for the largest N. This fact would have been exploited if it had been necessary to write a routine for solving simultaneous linear equations instead of simply calling one from the computer's library; but it was not worthwhile to write a routine specially because, as we shall now see, the reduction in the number of arithmetic operations performed during the whole calculation would have been too small to be useful. From [7] and [11], an ordinary routine for solving N linear equations in N unknowns requires about $N^3/3$ multiplications and the same number of additions, which for N = 28, 29, ..., 40 gives about 2.10^5 multiplications in total. Examination of eqn (38h) shows that evaluating each term in the sum giving a $u_{i,i'}$ requires two multiplications, two divisions and two subtractions, in addition to the time spent in determining the k_{Ai}, k_{Bj} and x_j; there are $N^2/2$ distinct $u_{i,i'}$, and each sum contains 1000 terms; so for N = 40 we must perform about 16.10^5 multiplications and the same number of divisions (which are at least as time-consuming). These operation counts are reduced in the author's program by evaluating and storing some of the quantities $(k_{Bj}^2/(k_{Bj}^2 - k_{Ai}^2))$ beforehand, but only by a factor of two (no greater saving than this is possible, because the j^{th} term of the sum defining a $u_{i,i'}$ depends on all three indices, i, i' and j). Consequently we could only reduce the running time of the program by a small fraction even if we reduced to negligible proportions the number of arithmetic operations required to solve the linear equations, so no special algorithms were used.

Finally, we consider the precision to which the calculations must be carried out. The description in Somlo's paper [6] and his program make it clear that the results in [6] were calculated using about 17 decimal places ("double-precision"). The author's experience, however, is that 11 decimal places (corresponding to the storing of a real variable in 48 bits) is sufficient; this is confirmed by the work reported in [9]. Single-precision arithmetic gave this number of places on the ICL-1906S computer used by the author, which saved substantially on time and storage.

20 ERRORS AND TOLERANCES IN PRACTICAL APPLICATION

The work described above was undertaken with the intention of using calculable coaxial-line discontinuities to construct calculable reflection and transmission standards. We therefore discuss briefly the errors likely to arise in such applications, considering (to fix ideas) the device shown in

Fig 4, which contains a section of oversize inner conductor and has normal-size inner on either side of this section. Devices of this kind are already in use as semi-calculable standards, both in the Microwave Standards Division of RSRE and at other standards laboratories, and it would be useful to make them fully calculable by doing away with the need for approximate theory of the discontinuity effects.

Assuming first that the dimensions of the device are exactly known, we remark that, since the errors in the capacitances at each of its two steps are a few parts in 10^4, and since the capacitances are responsible for about 10% of the total reflection (this was established by calculating the reflection with and without allowance for capacitance), the uncertainty in reflection coefficient due to direct error in the capacitances is a few parts in 10^5. Analysis over a wide frequency band is performed by interpolation between capacitance values calculated at "spot" frequencies, but as the frequency variation of capacitance is very small (1-2% between dc and 18 GHz in standard 7 mm line), the error due to inaccurate interpolation will be comparable with the errors in the "spot" values. Finally, the effect of mutual interference between the steps may be assessed as negligible. From [1], [2] and [13] it is known that the presence of a nearby reflector at a distance ℓ from a step modifies the reflection of an evanescent mode of decay constant γ by a factor varying between coth $(\frac{1}{2}\gamma\ell)$ and tanh $(\frac{1}{2}\gamma\ell)$ depending on the phase of the TEM excitation. For small reflection these quantities are both nearly 1, and their difference is $4e^{-\gamma\ell}$. If this difference is negligible for even the smallest significant γ (which may be obtained from the upper critical frequency discussed earlier) then interference can be ignored. As an example, the upper critical frequency for 7 mm line has already been determined (see section 6, "Generalised modes and critical frequencies") to be 75.1 GHz, so k' of eqn (2d) is $(2\pi/c) \times 75.1 \approx 1.5$ mm^{-1}. Up to 18 GHz the difference between k'^2 and $(k'^2 - k^2)$ is fairly small, so we may take $\gamma \sim k' \sim 1.5$ mm^{-1}. It is then clear that ℓ need only be a few millimetres to make $4e^{-\gamma\ell}$ an insignificant quantity; and the total error due to imperfections in calculating the capacitances will simply be the sum of the interpolation and direct calculation errors discussed above, about 1 in 10^4.

The situation is quite different when we come to consider the effect of imperfectly-known dimensions. This has been examined directly; four

devices of the type shown in Fig 4, based on standard 7 mm line and designed to have convenient maximum VSWRs, were analysed, and the calculations were then repeated with all the radii altered (in the directions of greatest effect) by 0.001 mm, which is the error of the measurements of radius as quoted by the RSRE metrology section. The general features of the frequency variation of the reflection coefficient of the device in Fig 4 are shown in Fig 5, and it is found that the predicted values of the reflection coefficient's maxima may change by as much as 0.002 for typical sizes of oversized inner conductor. These changes are small enough to give acceptably small uncertainties on the predicted values (less than 0.003 total uncertainty), but when they are compared with the uncertainties due to intrinsic error in the calculated capacitances, it is clear that the precision of the predictions is limited by the mechanical tolerances; the error of calculation is at least an order of magnitude smaller. A similar, but even stronger, conclusion is reached in the analysis of devices like the one in Fig 6, whose general behaviour is shown in Fig 7. The calculations for these were performed using a novel method of accurately analysing multi-port networks and cascaded two-ports, which will be described in a future paper. Comparisons of measurement and prediction for some of the devices represented in Figs 4 and 6, with a fuller error analysis, will also be presented at an early future date.

21 CONCLUSIONS

It has been shown that equivalent circuits for problems involving junctions of two generalised cylindrical waveguides can be found, and a method of calculating their component values has been fully described for the important special case of a coaxial line with a radial step on its inner conductor (including the degenerate form of this problem in which the inner conductor is truncated). The method has been programmed in Algol-68R on an ICL-1906S computer, for both the normal and degenerate forms of the problem, and gives results agreeing to a few parts in 10^4 with independent published work. The results are being applied to the development of calculable reflection and transmission standards for microwave measurements using coaxial lines; in this application it has been shown that the accuracy of prediction of performance is limited in practice only by mechanical tolerances.

REFERENCES

1 R F Harrington, "Time-harmonic electromagnetic fields", McGraw-Hill, New York (1961).

2 D S Jones, "The theory of electromagnetism", Pergamon, Oxford (1964).

3 D M Kerns and R W Beatty, "Basic theory of waveguide junctions and introductory microwave network analysis", International Series of Monographs in Electromagnetic Waves (vol 13), Pergamon, New York (1967).

4 G N Watson, "A treatise on the theory of Bessel functions" (2nd edition), Cambridge University Press (1944).

5 R Courant and D Hilbert, "Methods of mathematical physics" (vol 1) (translated from the German text), Interscience, New York (1953).

6 P Somlo, "The computation of coaxial-line step capacitances", Trans IEEE vol MTT-15 (1967), pp 48-53.

7 D K Faddeev and V N Faddeeva, "Computational methods of linear algebra" (translated from the Russian text), W H Freeman and Co, San Francisco (1963).

8 J R Whinnery, H W Jamieson and T E Robbins, "Coaxial-line discontinuities", Proc IRE vol 32 (1944), pp 695-709.

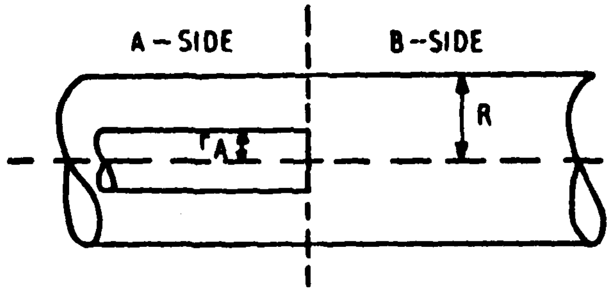
9 B Bianco, A Corana, L Gogioso, S Ridella and M Parodi, "Open-circuited coaxial lines as standards for microwave measurements", Electron Lett vol 16 no 10 (9/5/1980), pp 373-374.

10 D Shanks, "Non-linear transformations of divergent and slowly-convergent sequences", J Math Phys vol 34 (1955), pp 1-42.

11 J H Wilkinson, "Error analysis of direct methods of matrix inversion", J Assoc Comput Mach vol 8 (1961), pp 281-330.

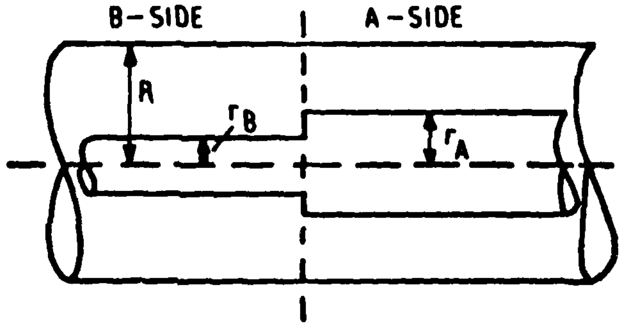
12 "Collected algorithms from ACM" (vol 1), Association for Computing Machinery, New York (1980). (Algorithm 215 originally appeared in Comms ACM vol 6 (1963), p 662).

13 R E Collin, "Field theory of guided waves", International Series in Pure and Applied Physics, McGraw-Hill, New York (1960).



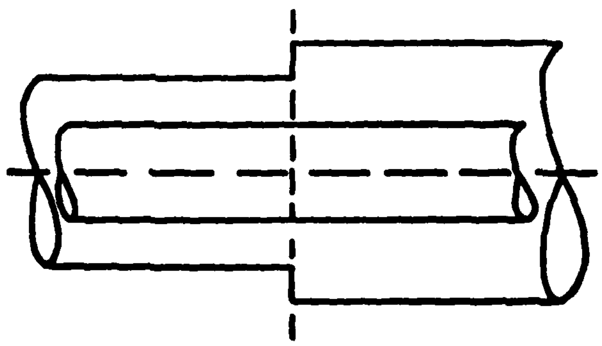
(TRUNCATED INNER)

FIG.1A



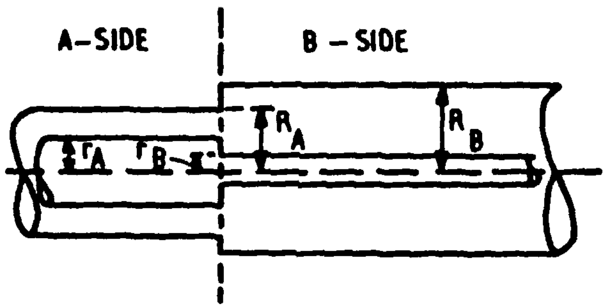
(STEPPED INNER)

FIG.1B



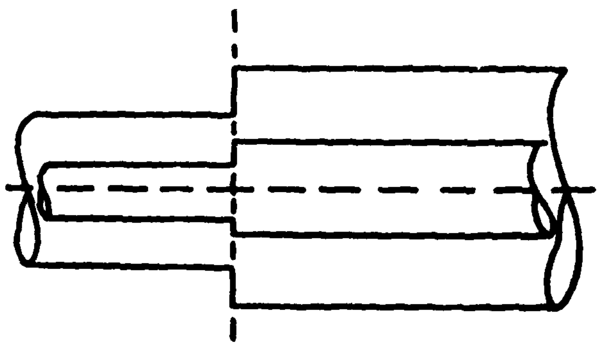
(STEPPED OUTER)

FIG.1C



(DOUBLE STEP)

FIG. 1D



(DOUBLE STEP)

FIG. 1E

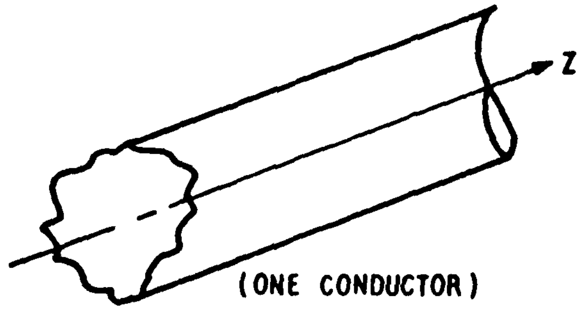


FIG. 2A

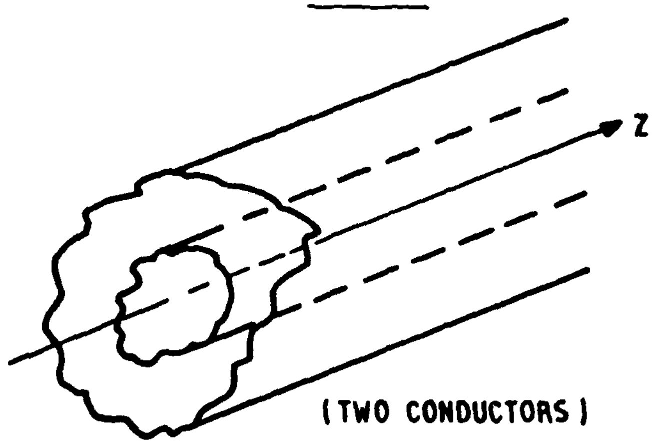


FIG. 2B

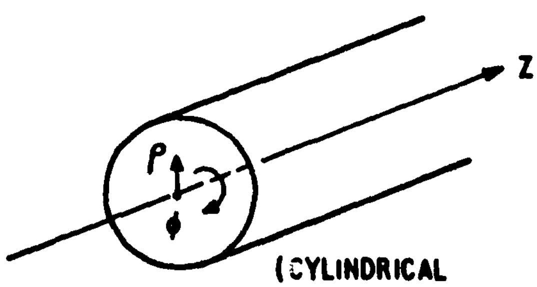


FIG. 3

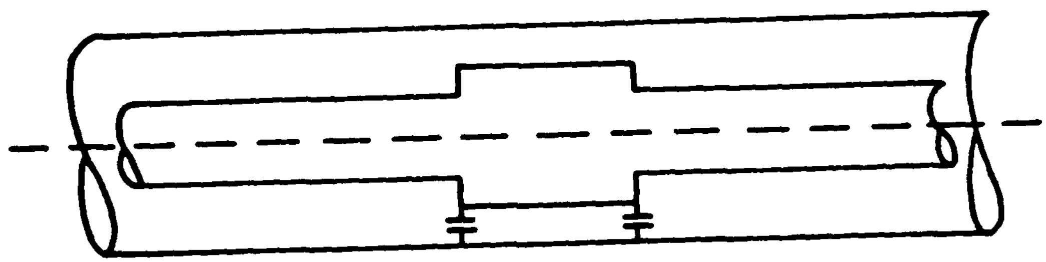


FIG. 4

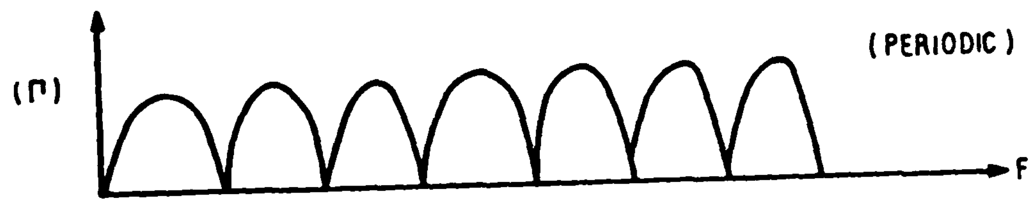


FIG.5

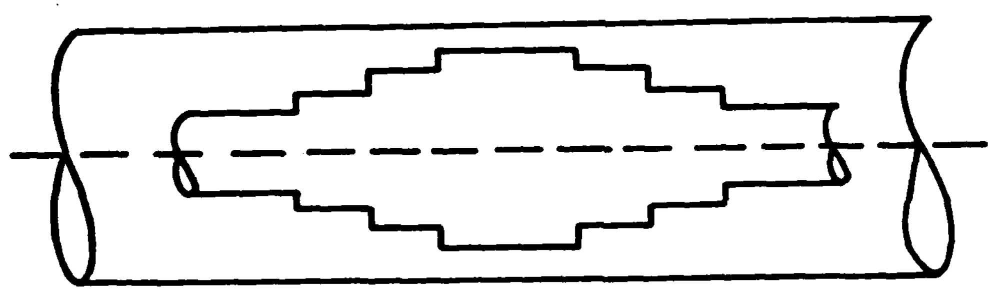


FIG.6

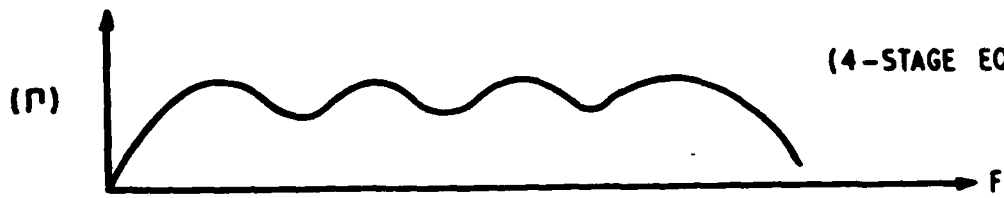


FIG.7

EXPERIMENTIAL STEPPED-INNER-CONDUCTOR REFLECTION STANDARD DESIGNS

EN

DAT

FILM

4-8

DTI